Characterizing dose-response model uncertainty using model averaging.

Matt Wheeler
NIOSH/EID/REB

National Institute for Occupational Safety and Health
NIOSH

CDC
Workplace Safety and Health

Characterizing dose-response model uncertainty using model averaging.

Acknowledgements

- A. John Bailer

  Miami University/NIOSH

Outline

- Introduction/Motivation
- Model Averaging
- Simulation
- Conclusions/Future research

Characterizing dose-response model uncertainty using model averaging.

Introduction

- Given an experimental hazard data, we are frequently concerned with estimating a level of exposure, that corresponds to level of risk for the hazard of interest.
- This value, called the benchmark dose, is estimated based upon a chosen regression model.
- Multiple models are frequently available and often describe the data “equally.”
- Even though these models describe the data similarly the models often characterize the risk at low levels of exposure differently.
- Yet decisions are made with one model in mind.

Characterizing dose-response model uncertainty using model averaging.

Introduction

- Consider the problem of estimating a benchmark dose (BMD) from dichotomous dose response data.
- Here we seek to estimate the BMD from a “plausible” model, given experimental data.
- In these experiments:
  - Animals are exposed to some potential hazard.
  - The adverse response is assumed to be distributed binomially.
  - Risk (i.e, probability of adverse response) is estimated using regression modeling.
  - Multiple dose-response models can be used to estimate risk.

Characterizing dose-response model uncertainty using model averaging.

Introduction

Common Dose-Response Models Used:

- logistic model: $$\pi_1(d) = \frac{1}{1+\exp[-(\alpha+\beta d)]} \quad (1)$$
- log-logistic model: $$\pi_2(d) = \gamma + \frac{(1-\gamma)}{1+\exp[-(\alpha+\beta\ln(d))]} \quad (2)$$
- gamma: $$\pi_3(d) = \gamma + (1-\gamma)\frac{1}{\Gamma(\alpha)}\int_0^{\beta d} t^{\alpha-1}e^{-t}dt \quad (3)$$
- multistage $$\pi_4(d) = \gamma + (1-\gamma)\left[1-\exp(-\theta_1 d - \theta_2 d^2 ...)\right] \quad (4)$$
- probit $$\pi_5(d) = \Phi(a+\beta d) \quad (5)$$

Characterizing dose-response model uncertainty using model averaging.

Introduction

- log-probit $$\pi_6(d) = \gamma + (1-\gamma)\Phi\left[a + \beta \ln(d)\right] \quad (6)$$
- quantal-linear $$\pi_7(d) = \gamma + (1-\gamma)\left[1 - \exp(-\beta d)\right] \quad (7)$$
- quantal-quadratic $$\pi_8(d) = \gamma + (1-\gamma)\left[1 - \exp(-\beta d^2)\right] \quad (8)$$
- Weibull $$\pi_9(d) = \gamma + (1-\gamma)\left[1 - \exp(-\beta d^{\alpha})\right] \quad (9)$$

where $\Gamma(\alpha)$= gamma function evaluated at α, for $\Phi(x)$ = CDF N(0,1) and $\pi_i = \gamma$ when d_i=0 for models (2) and (7).

Characterizing dose-response model uncertainty using model averaging.

Introduction

- Given data (in absence of mechanistic information), a typical analyst will:
  - Estimate the regression coefficients for models (1)-(9).
  - Estimate the BMD/BMDL given the model.
  - Pick the “best model.”
- As uncertainty results from one given model, a different approach may be helpful.

Characterizing dose-response model uncertainty using model averaging.

Model Averaging

- A better way would be to find an adequate way to combine all estimates, and thus describe/account for model uncertainty.
- Model Averaging (MA) is a method that may satisfactorily account for model uncertainty.
- Instead of focusing on a single model it allows researchers to focus on “plausible behavior.”

Characterizing dose-response model uncertainty using model averaging.

Model Averaging

- Given the fits of models (1)-(9) MA:
  - Calculates the dose-response based upon a weighted average of dose-responses Raftery et al. (1997), Buckland et al. (1997)
  - Estimates the MA dose-response curve as:

$$\pi_{\mathrm{MA}}(d) = \sum_{i=1}^{K} \pi_i(\boldsymbol{\theta}, d) \cdot w_i$$

  - Weights are formed as:

$$w_j = \frac{\exp(-I_j / 2)}{\sum_{i=1}^{K} \exp(-I_i / 2)}$$

  - Where I_i=AIC, I_i=KIC , or I_i=BIC. Other weights are possible.

Characterizing dose-response model uncertainty using model averaging.

Model Averaging

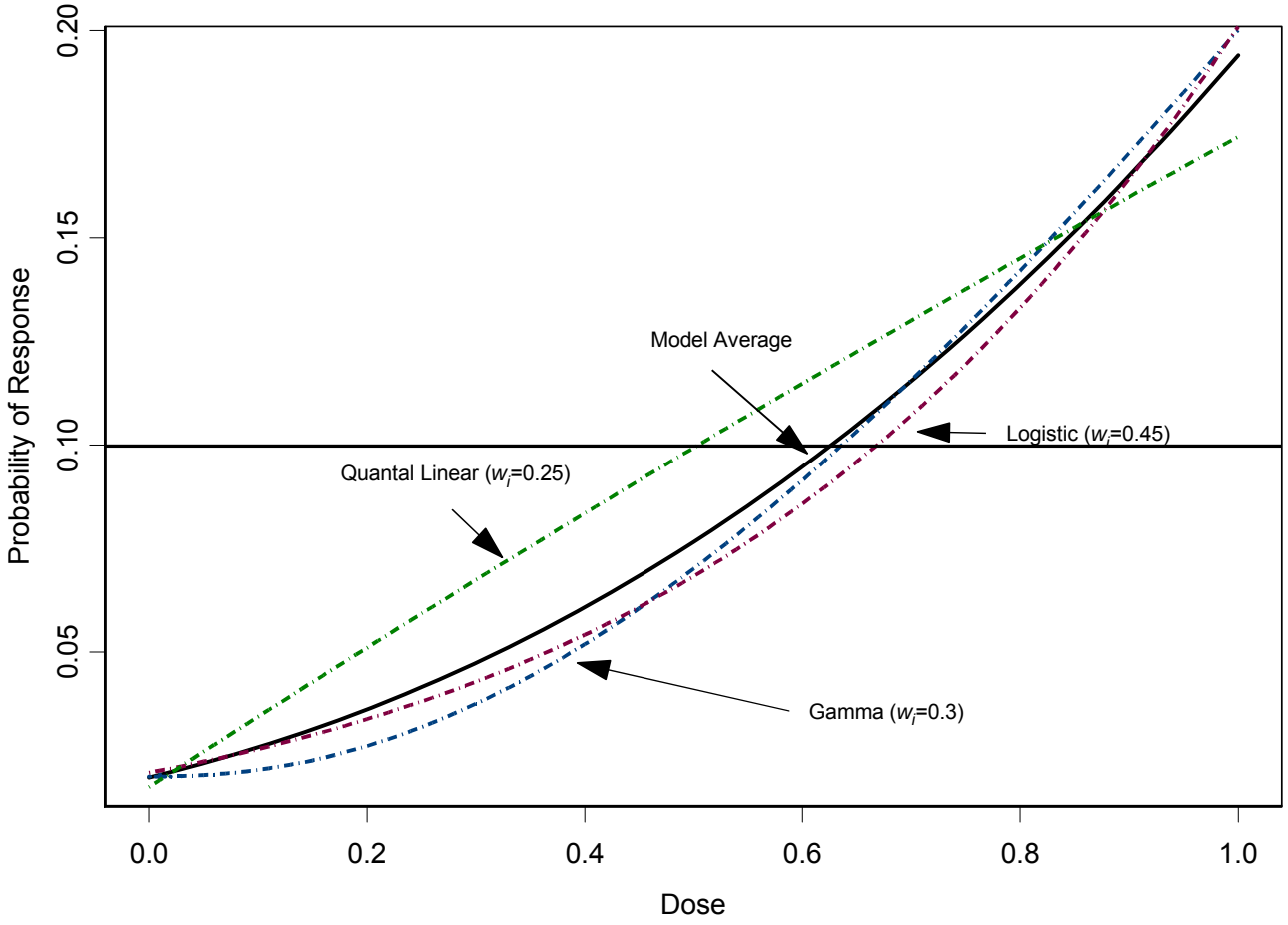
- Given this "Average-model," the benchmark dose is then computed by finding the dose that satisfies the equation

$$BMR = [\pi_{MA}(d)_i - \pi_{MA}(0)]/[1 - \pi_{MA}(0)].$$

- BMR typically set at values of 1, 5, and 10%.
- The BMDL is computed through a parametric bootstrap. Here the 5th percentile of the bootstrap distribution is used to compute the 95% lower tailed confidence limit estimate on the BMD.

Characterizing dose-response model uncertainty using model averaging.

Model Averaging



Characterizing dose-response model uncertainty using model averaging.

Model Averaging

- MA seems like a good idea, however we need to know if it works well in practice.
- A simulation study was conducted investigating the behavior of MA.
- 54 true model conditions, using models (1) – (9), were used in the simulation.
- Full study described in Wheeler and Bailer (Risk Analysis, *In Press*)

Characterizing dose-response model uncertainty using model averaging.

Simulation

- The simulation proceeded by generating hypothetical toxicology experiments with response probability $\pi(d)$.
- With $\pi(d)$ specified by one of the 54 true dose-response curves.
- These experiments consisted of 4 dose group design with doses of 0, 0.25, 0.50, and 1.0.
- n=50 for all dose groups.
- 2000 experiments were generated per true dose-response curve.
- Bias as well as coverage [i.e., $\Pr(\text{BMDL} \leq \text{BMD}_{\text{true}})$] was estimated.
- Coverage is reported here.

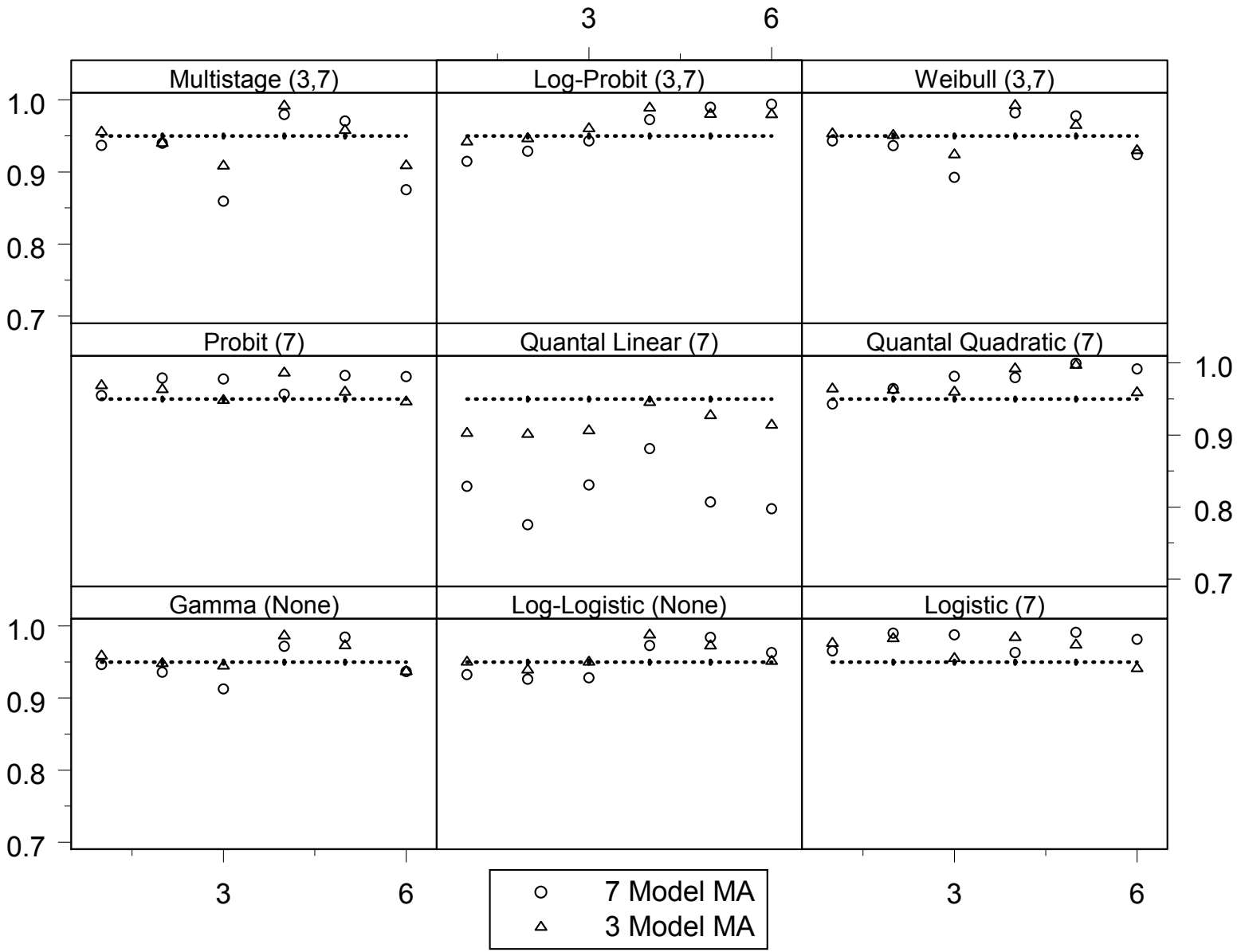
Characterizing dose-response model uncertainty using model averaging.

Simulation

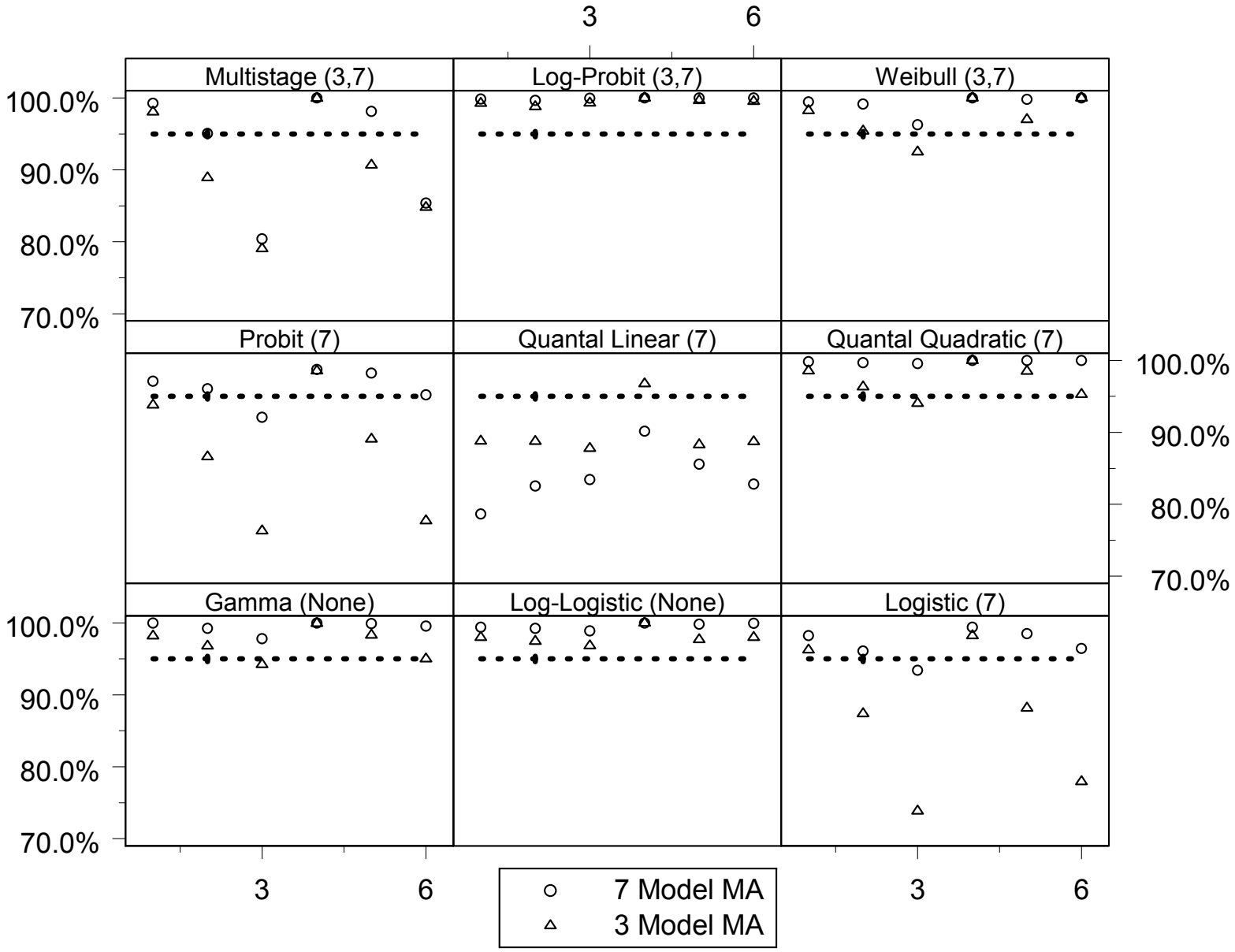
- In each experiment the “average-model” BMD as well as the BMDL was estimated.
- BMRs of 1% and 10% were used to estimate the BMD.
- Two model spaces for averaging were considered.
  - One space consisted of three flexible models: the multistage, Weibull and the log-probit model.
  - The second space had seven models that added the probit, logistic, quantal-linear, and quantal-quadratic to the three model space.

- Coverage probability [i.e., $Pr(BMDL \leq BMD_{true})$] was estimated across 2000 simulations.
- The nominal coverage level was 95%.
- The simulation took approximately 1 CPU year of computation.

Coverage BMR = 10%



Coverage BMR = 1%

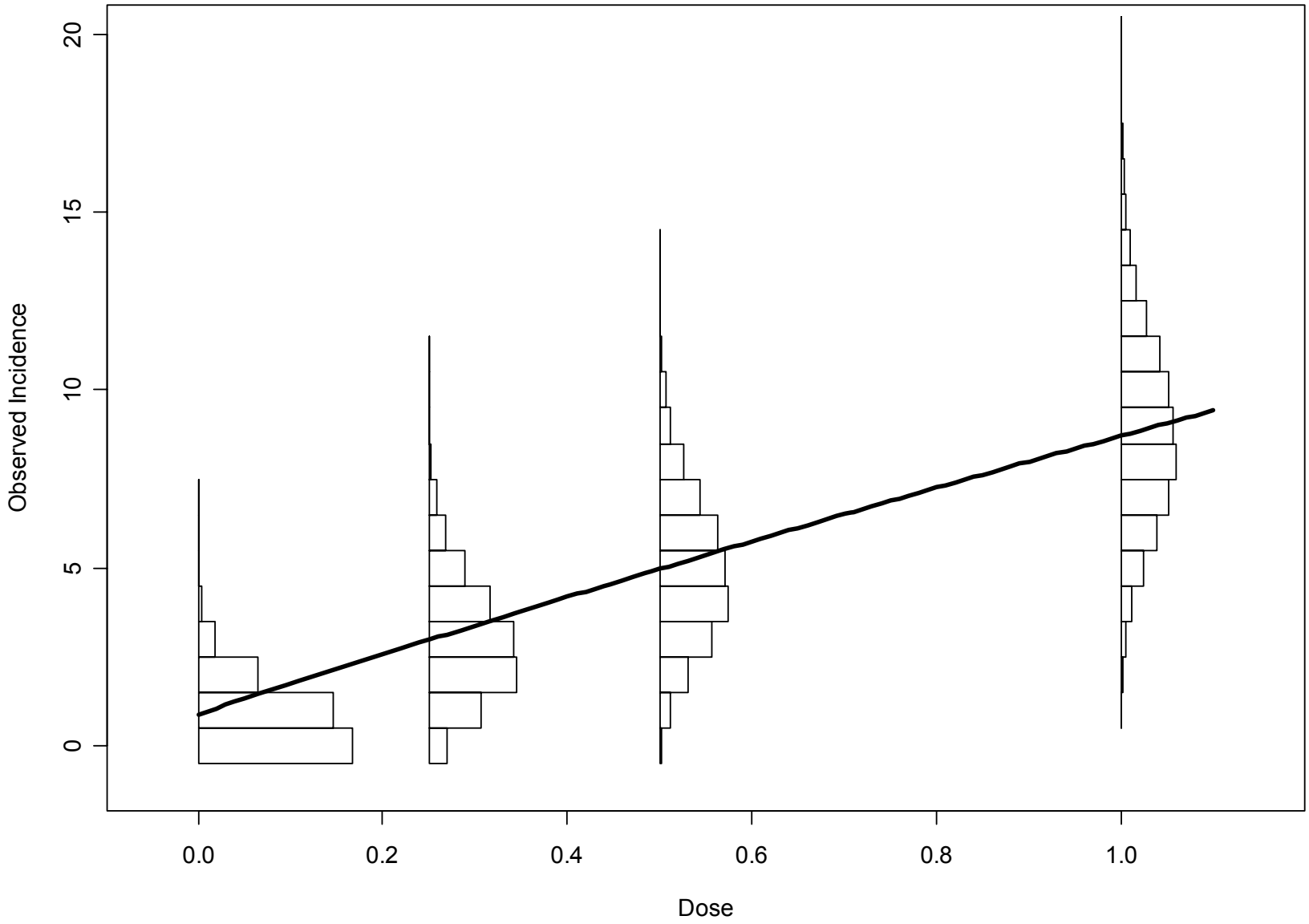


Characterizing dose-response model uncertainty using model averaging.

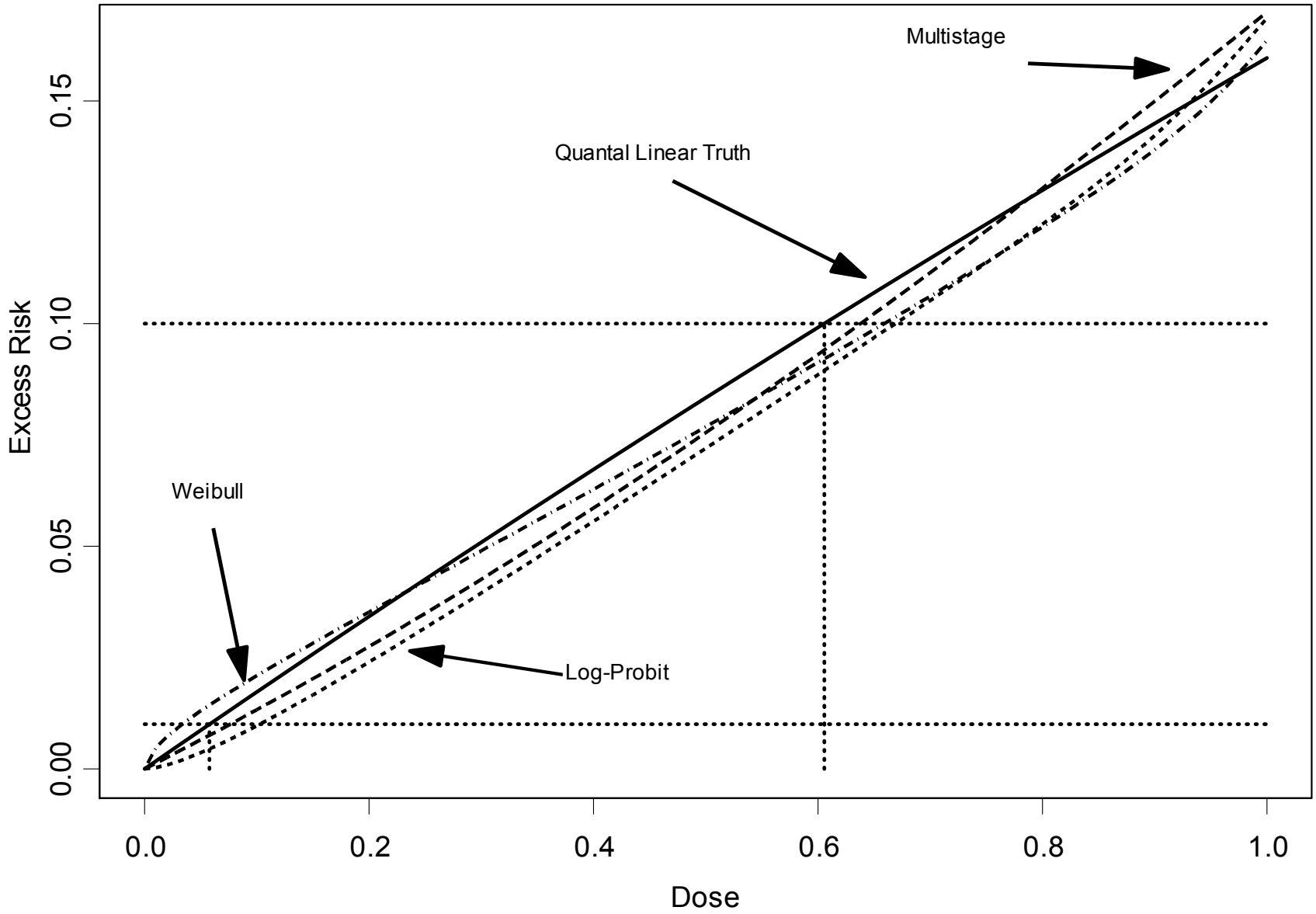
Simulation

- Nominal coverage is reached for most simulation conditions.
- MA fails to reach nominal coverage in the quantal-linear and similar cases.
- It is important to understand why the BMD is mischaracterized in the quantal linear case.
- We study this through investigating the sampling distribution.

Sampling distribution for the quantal-linear model



Average fit for 3-model MA models



Characterizing dose-response model uncertainty using model averaging.

Simulation

- The flexibility of the models combined with the sampling distribution introduces bias into the estimation of the dose-response curve.
- The bias carries through in BMD estimation.
- This also may be the cause of the conservative behavior (i.e. coverage > 99%) seen in the quantal-quadratic case.

Characterizing dose-response model uncertainty using model averaging.

Simulation

- Improved coverage can be obtained using BCa intervals.
- Other results suggest that MA is superior to picking the best model.
- The results show MA is not a panacea, it is however a step in the right direction.

Characterizing dose-response model uncertainty using model averaging.

Simulation

- The results are promising but implementation of this approach is difficult.
- The simulation code has been repackaged to allow users to implement dichotomous dose-response model averaging.
- This is done in a simple MS Windows command prompt program.

Characterizing dose-response model uncertainty using model averaging.

Conclusions/Future research

- As mentioned before model averaging is not a panacea.
- As such it does not:
  - Relieve scientists from using their expert judgment.
  - Give automatic license to produce a low dose extrapolations.
  - Remove the need for adequate individual model fit diagnostics.
  - Remove all model uncertainty from the analysis.

Characterizing dose-response model uncertainty using model averaging.

Conclusions/Future research

- It does:
  - Reframe the debate of model choice.
  - Produces relatively stable central estimates often independent of a given model being included in the average.
  - Point direction to future research.

Characterizing dose-response model uncertainty using model averaging.

Conclusions/Future research

- Future research:
  - Continuous and count data MA software development.
  - Extensive Study of a proper suite of models to use in MA.
  - Study of experimental designs that might optimize MA performance in terms of estimation and lower bound calculation.

Thank You